

Exploring Medicaid Data for Research: Potentials and Pitfalls

For CERCIT Workshop
(7-22-2011)
By Xianglin L. Du, Ph.D.
UT-School of Public Health
UTHealth in Houston

Outline of the Talk

- Medicaid eligibility / brief overview
- Medicaid Data processing brief background
- Medicaid MAX Data Files and Contents
- Medicaid Data quality and validation studies - reviews
- Potentials/Advantages of Medicaid Data
- Pitfalls/Limitations of Medicaid Data

Medicaid Data – brief review

- See previous helpful slides/notes on Medicaid data overview by TCR team (Melanie William & Cheryl Bowcock) on 3/25/11 seminar.
- Medicaid is health insurance that helps many people who can't afford medical care pay for some or all of their medical bills.
- Medicaid is available only to people with limited income.

“Good health is important to everyone. If you can't afford to pay for medical care right now, Medicaid can make it possible for you to get the care that you need so that you can get healthy and stay healthy.”

Medicaid Eligibility

Mandatory Eligibility Groups (copied from <https://www.cms.gov/>)

- Limited income families with children, as described in Section 1931 of the Social Security Act, who meet certain of the eligibility requirements in the state's Aid to Families with Dependent Children (AFDC) plan in effect on July 16, 1996;
- Supplemental Security Income (SSI) recipients (or in States using more restrictive criteria—aged, blind, and disabled individuals who meet criteria which are more restrictive than those of the SSI program and which were in place in the State's approved Medicaid plan as of January 1, 1972);
- Infants born to Medicaid-eligible pregnant women. Medicaid eligibility must continue throughout the first year of life so long as the infant remains in the mother's household and she remains eligible, or would be eligible if she were still pregnant;
- Children under age 6 and pregnant women whose family income is at or below 133% of the Federal poverty level. (The minimum mandatory income level for pregnant women and infants in certain States may be higher than 133% percent, if as of certain dates the State had established a higher percentage for covering those groups.) States are required to extend Medicaid eligibility until age 19 to all children born after September 30, 1983 (or such earlier date as the State may choose) in families with incomes at or below the Federal poverty level. Once eligibility is established, pregnant women remain eligible for Medicaid through the end of the calendar month in which the 60th day after the end of the pregnancy falls, regardless of any change in family income. States are not required to have a resource test for these poverty level related groups. However, any resource test imposed can be no more restrictive than that of the AFDC program for infants and children and the SSI program for pregnant women;
- Recipients of adoption assistance and foster care under Title IV-E of the Social Security Act;
- Certain people with Medicare; and
- Special protected groups who may keep Medicaid for a period of time. Some examples include people who lose SSI payments due to earnings from work or increased Social Security benefits; and families who are provided 6 to 12 months of Medicaid coverage following loss of eligibility under Section 1931 due to earnings, or 4 months of Medicaid coverage following loss of eligibility under Section 1931 due to an increase in child or spousal support.

Medicaid Eligibility (cont'd)

Optional Eligibility Groups (copied from <https://www.cms.gov/>)

- Infants up to age one and pregnant women not covered under the mandatory rules whose family income is below 185% of the Federal poverty level (the percentage to be set by each state);
- Optional targeted low-income children;
- Certain aged, blind, or disabled adults who have incomes above those requiring mandatory coverage, but below the Federal poverty level;
- Children under age 21 who meet income and resources requirements for the Aid to Families with Dependent Children (AFDC), but who otherwise are not eligible for the AFDC;
- Institutionalized individuals with limited income and resources;
- Persons who would be eligible if institutionalized but are receiving care under home and community-based services waivers;
- Recipients of state supplementary payments;
- Tuberculosis-infected persons who would be financially eligible for Medicaid at the Supplemental Security Income (SSI) level (only for TB-related ambulatory services and TB drugs); and
- Low-income, uninsured women screened and diagnosed through a Center's for Disease Control (CDC) Breast and Cervical Cancer Early Detection Program (NBCCEDP) and determined to be in need of treatment for breast or cervical cancer. (Additional information may be found under Downloads and Related Links Inside CMS listed below).

Medicaid Data Background

- (based on CMS-ORDI's maxintro by Susan Radke and David Baugh, 5/6/2008)
- **Medicaid Tape-to-Tape Project**
 - 1980 to early 1990s
 - Five states
 - Prototype for later data collection
- **State Medicaid Research Files (SMRF)**
 - From 1987 to 1998 Voluntary participation by states (5-38 states)
- **Medicaid Analytic eXtract (MAX)**
 - Beginning in 1999 all states were mandated to submit Medicaid data to CMS via the Medicaid Statistical Information System (MSIS).

MAX Data Purpose

- (based on maxintro in <http://www.ncvhs.hhs.gov/081119p8.pdf> by David Baugh, 11/19/2008)
- **Purpose**
 - Produce data to support research and policy analysis on Medicaid and SCHIP (State Children's Health Insurance Program) populations.
 - MAX is needed because source data are not organized to support research
- MAX is derived from the Medicaid Statistical Information System (MSIS) for all states & D.C.
- **Major Differences between MAX and MSIS:**
For Medicaid services, MAX combines initial claims, voids and other adjustments to create a "final action event"

Whose data is available in Medicaid?

- The CMS Medicaid enrollment and claims data contains information for Medicaid eligibles who actually do enroll in their state's Medicaid program. CMS Medicaid files cannot, therefore, be used to study individuals who are eligible but not enrolled.
- Q: what is the study population?

Eligible vs. Enrolled vs. Enrolled+Used Services

Medicaid MAX Data Files

- **Personal Summary File (one record per case):**
- **Claims (Services) Files (multiple records per case):**
 - Inpatient Hospital (IP)
 - Long Term Care (LT)
 - Other Services (OT)
 - Prescription Drug (RX)

MAX Data Files (cont'd)

What types of data elements are in each of 5 files?

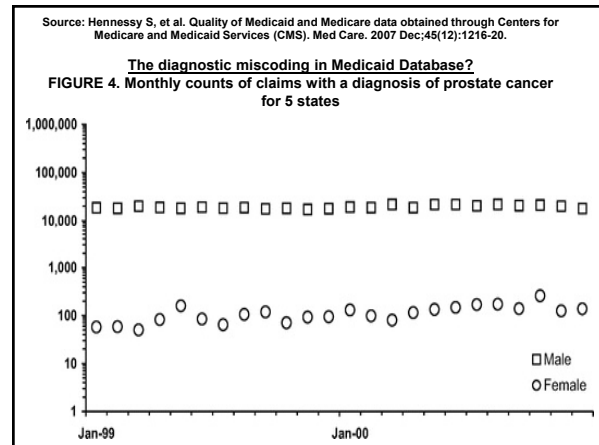
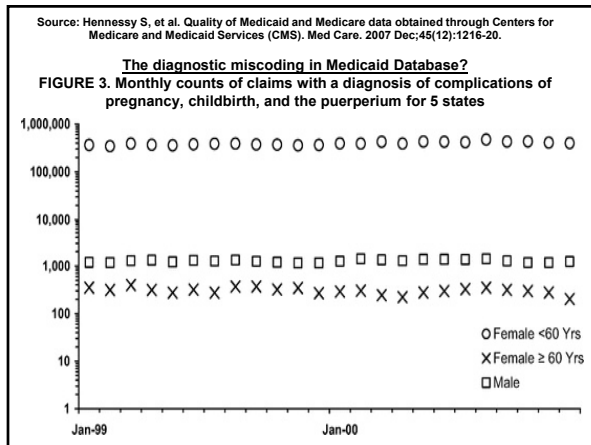
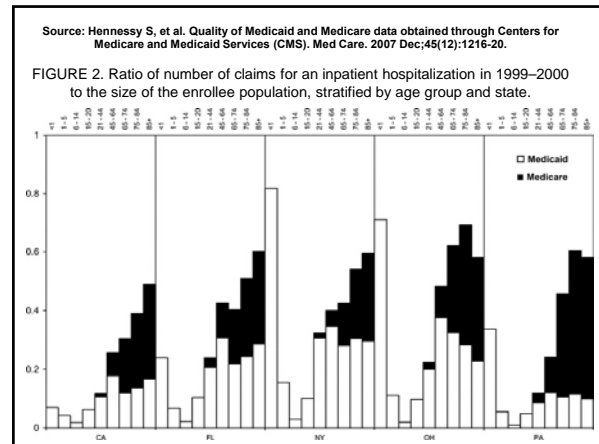
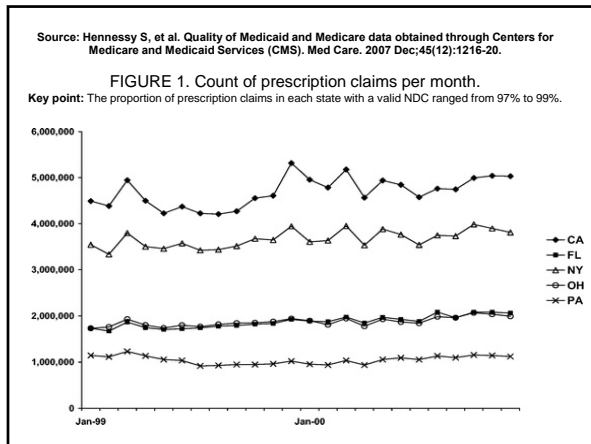
- **Personal Summary File:**
one record for every individual enrolled for at least one day during the year. The file contains demographic data (e.g. date of birth, gender, race), basis of eligibility, maintenance assistance status, monthly enrollment status, and a utilization summary.
- **Inpatient File:**
complete stay records for enrollees who used inpatient services including 10 diagnoses, 7 procedures, discharge status, length of stay, and payment amount.
- **Long Term Care File:**
claims for long term care services provided by Skilled Nursing Facilities (SNFs), Intermediate Care Facilities (ICFs), and independent psychiatric facilities. Fields include facility type, 5 diagnoses, dates of service, and discharge status.
- **Drug File:**
final action paid drug which beginning with 1996 data required an NDC.
- **Other Therapy File:**
claim records for all non-institutional Medicaid services, including physician services, lab/X-ray, clinic services and premium payments. As appropriate the claims include 2 diagnoses, 1 procedure, and date of service.

Medicaid Data Quality/Validation-review

- **Data quality and validation studies?**
 - 5 states (California, Florida, New York, Ohio, Pennsylvania)
 - California TR
 - New York study
 - Washington
 - Michigan
 - North Carolina
 - Texas
 - Many other states with non-tumor conditions Medicaid data validation.

Quality of Medicaid Data (in 5 States: CA, FL, NY, OH, PA)

- Source: Hennessy S, et al. Quality of Medicaid and Medicare data obtained through Centers for Medicare and Medicaid Services (CMS). Med Care. 2007 Dec;45(12):1216-20.
 - **OBJECTIVES:** To perform exploratory analyses of Medicaid and Medicare data obtained through CMS.
 - **RESEARCH DESIGN:** Obtained data from 1999 to 2000 for 5 large Medicaid programs (California, Florida, New York, Ohio, and Pennsylvania), together with corresponding Medicare data for dual eligible.
 - **RESULTS:** by 1) The number of prescription claims per month –stable or not; 2) hospitalizations rates ni Medicaid vs. supplemental Medicare claims; 3). **Diagnostic mis-coding in Medicaid** (diagnoses of complications of pregnancy and childbirth were uncommon in men and in women older than 60 years, and prostate cancer diagnoses were uncommon in women).



**Quality of Medicaid Data (cont'd)
(in 5 States: CA, FL, NY, OH, PA)**

- Source: Hennessy S, et al. Quality of Medicaid and Medicare data obtained through Centers for Medicare and Medicaid Services (CMS). Med Care. 2007 Dec;45(12):1216-20.
 - RESULTS: The number of prescription claims per month was very stable. Because of cocoverage by Medicare, Medicaid data seemed to miss a substantial number of hospitalizations in those age "≥65 years as well as 45-64 years".
 - "The diagnostic miscoding was infrequent: diagnoses of complications of pregnancy and childbirth were uncommon in men and in women older than 60 years, and prostate cancer diagnoses were uncommon in women."
 - CONCLUSIONS: In contrast to Medicaid data previously obtained from a commercial vendor, we found no evidence that prescription Medicaid data from CMS were incomplete. Researchers using Medicaid data to study hospital outcomes should obtain supplemental Medicare data on dual eligibles for studies of persons aged 45 years and older.

**Medicaid vs. Tumor Registry
(in California)**

- Source: Schrag D, et al. Linking tumor registry and Medicaid claims to evaluate cancer care delivery. Health Care Financ Rev. 2009;30(4):61-73.
 - compared cancer treatment from Medicaid claims with CCR (California Cancer Registry) data, using 1988-2000 cases matched with 1997-1998 Medicaid enrollment data.
 - Medicaid claims corroborated diagnoses for 73% of breast and 68% of colorectal cancers in CCR.
 - Medicaid claims confirmed surgery for 67% of CCR's breast cancers.

Source: Schrag D, et al. Linking tumor registry and Medicaid claims to evaluate cancer care delivery. Health Care Financ Rev. 2009;30(4):61-73.

Table 1
Number of 1998 Incident Primary Cancers, by Site Reported to the California Cancer Registry (CCR) and the Proportion of Cancer Patients age 18-64 Enrolled in the California Medicaid Program

Tumor Site	Incident Cancer Cases from the California Cancer Registry		Cases in both the California Cancer Registry and California Medicaid	
	All 1998 Cases	1998 Cases Age 18-24	% Cases Age 18-24	Cases Age 18-64 and in Medicaid for >=1 month in 1998
Breast	20,864	11,382	54.60%	1,014
Lung	17,004	5,203	30.60%	997
Colorectal	10,254	2,796	27.30%	557
Cervical	1,690	1,350	79.90%	287
Prostate	19,001	5,605	29.50%	268
Testis	937	877	93.60%	78
Uterine	3,587	1,722	48%	176
Bladder	5,452	1,445	26.50%	113
Hepatoma	1,570	856	53.30%	288
Stomach	2,592	822	32.50%	197
Ovarian	2,626	1,468	55.90%	188
Kidney	2,914	1,358	46.60%	173
Brain	1,922	1,246	64.80%	206

Source: Schrag D, et al. Linking tumor registry and Medicaid claims to evaluate cancer care delivery. Health Care Financ Rev. 2009;30(4):61-73.

Table 3
Cancer Diagnoses and Procedures in California Medicaid Claims for Patients with Cancer Diagnoses Recorded in CCR. Patients Were Enrolled in Medicaid for at Least Some Portion of 1998 and Age 18-64 at Diagnosis.

Cancer Type	Diagnosed At Any Time During 1998		Diagnosed in the First Half of 1998	
	Number of CCR-Enrolled with Cancer	Number (%) with a Corroborating Diagnosis Code Recorded in 1998 Medicaid Claims	Number of CCR-Enrolled with Cancer	Number (%) with a Corroborating Diagnosis Code Recorded in 1998 Medicaid Claims Files
Breast	1,014	738 (73%)	549	400 (74%)
Lung	997	751 (75%)	523	416 (80%)
Colorectal	557	378 (68%)	274	189 (69%)
Cervical	287	207 (72%)	146	111 (76%)
Prostate	268	178 (66%)	127	93 (73%)
Testis	78	58 (77%)	37	28 (75%)
Uterine	176	121 (69%)	86	62 (72%)
Bladder	113	86 (76%)	53	42 (79%)
Liver	288	180 (63%)	151	95 (63%)
Stomach	197	137 (69%)	91	70 (77%)
Ovarian	188	130 (69%)	100	74 (76%)
Kidney	173	118 (68%)	93	70 (75%)
Brain	206	128 (65%)	98	64 (65%)

Source: Schrag D, et al. Linking tumor registry and Medicaid claims to evaluate cancer care delivery. Health Care Financ Rev. 2009;30(4):61-73.

Table 4
The Sensitivity of California Medicaid Claims for Identifying Breast Cancer Surgeries among Women with Primary Breast Cancer Diagnoses and Breast Cancer Surgery in the First Half of 1998 Based on California Cancer Registry (CCR) Records

Cohort Definition	Denominator	YES, Corroborating Surgery Claim Identified		NO, Corroborating Surgery Claim Not Identified	
		N	%	N	%
How often was there a claim for breast surgery in the Medicaid records of women who had breast cancer diagnoses and breast surgery recorded in CCR?	467	331	71%	136	29%
Among the subgroup enrolled in a managed care plan?	240	156	65%	84	35%
Among the subgroup not enrolled in a managed care plan?	227	175	77%	52	23%
How often were there claims for breast surgery in 1998 Medicaid records of women continuously enrolled in Medicaid during 1998?	239	174	73%	65	27%
How often were there claims for breast surgery in 1998 Medicaid files for women enrolled for at least 1 but fewer than 12 months of 1998 in a Medicaid plan?	228	153	68%	75	32%
How often were there any inpatient Medicaid claims recorded for women who had primary breast cancer surgery recorded by the CCR, but no claim for breast surgery identifiable in Medicaid files?	136	45	33%	91	67%

Medicaid vs. Tumor Registry (in California) (cont'd)

- Source: Schrag D, et al. Linking tumor registry and Medicaid claims to evaluate cancer care delivery. Health Care Financ Rev. 2009;30(4):61-73.
- ...
 - Medicaid claims corroborated diagnoses for 73% of breast and 68% of colorectal cancers in CCR.
 - Medicaid claims confirmed surgery for 67% of CCR's breast cancers.
 - **Conclusion:** ... Medicaid claims have moderate sensitivity for identifying cancer diagnoses and surgery. Linked registry-Medicaid data can identify indigent patients and the timing of Medicaid coverage.

Medicaid vs. Tumor Registry (in New York)

- Source: Boscoe FP, et al. Building capacity to assess cancer care in the Medicaid population in New York state. Health Serv Res. 2011 Jun;46(3):805-20.
 - **Data Sources.** Over 500,000 cancer patients diagnosed in 2002-2006 reported to the New York State Cancer Registry were linked with New York State Medicaid enrollment and claims records.
 - **Principal Findings.**
 - One-quarter of cancer patients were enrolled in Medicaid at or near the time of cancer diagnosis.
 - The Medicaid cohort was younger, more likely to be an ethnic minority, foreign born, never married, live in either an inner-city or remote rural area, and have a higher stage at diagnosis.

Source: Boscoe FP, et al. Building capacity to assess cancer care in the Medicaid population in New York state. Health Serv Res. 2011 Jun;46(3):805-20.

Table 2. Percentage of Cancer Patients Enrolled in Medicaid,* Adults and Elderly, in Descending Order by Cancer Site, New York State, 2002-2006

Site	% Enrolled	
	Adults (18-64)	Elderly (65+)
All sites combined	27.3	23.2
Liver	52.2	33.6
Cervix	50.8	48.9
Anal	44.4	31.7
Larynx	42.9	27.9
Stomach	41.7	33.0
Esophagus	38.7	25.4
Oral	38.5	27.0
Hodgkin lymphoma	37.0	26.5
Lung	35.5	22.8
Multiple myeloma	35.5	28.8
Brain	32.8	29.8
Non-Hodgkin lymphoma	32.0	18.8
Leukemia	31.8	20.8
Pancreas	30.5	22.4
Rectum	30.0	27.7
Colon	29.7	27.0
Ovary	27.0	24.5
Kidney	26.5	21.4
Testis	25.8	25.4
Uterus	23.4	27.2
Female breast	22.2	24.4
Thyroid	22.1	23.6
Bladder	19.6	18.5
Prostate	17.0	18.5
Melanoma	9.9	9.9

* "Enrolled in Medicaid" means ever enrolled in Medicaid during the 2001-2009 period; it does not mean enrolled at the time of diagnosis.
 *† Sites with values that are higher than for all sites combined are overrepresented in the Medicaid population; sites with values lower than this figure are underrepresented.

Source: Boscoe FP, et al. Building capacity to assess cancer care in the Medicaid population in New York state. Health Serv Res. 2011 Jun;46(3):805-20.

Table 3. Cancer Stage Distribution by Medicaid Enrollment Status,* New York State, 2002–2006

Stage [‡]	Lung		Colorectal		Female Breast	
	Medicaid	Non-Medicaid	Medicaid	Non-Medicaid	Medicaid	Non-Medicaid
In situ	—	—	7.3	8.8	15.4	23.0
Local	15.3	18.7	27.7	32.8	42.0	47.0
Regional	22.0	22.7	35.4	33.4	27.5	20.5
Distant	45.3	42.8	16.5	14.8	6.6	3.3
Unknown	17.3	15.7	13.1	10.2	8.4	6.2

* Enrolled in Medicaid at any time during the 2001–2008 period.
[‡] SEER Summary Stage 2000.

Medicaid vs. Tumor Registry (in New York) (cont'd)

• Source: Boscoe FP, et al. Building capacity to assess cancer care in the Medicaid population in New York state. Health Serv Res. 2011 Jun;46(3):805-20.

– **Conclusions.** The linked dataset will permit detailed analysis of cancer treatment and cancer treatment disparities among historically understudied groups. The linkage has also resulted in improvements in Cancer Registry quality through the identification of errors and missing values.

Medicaid-Medicare Data (in Michigan)

• Source: Bradley CJ, et al. Adjuvant chemotherapy after resection in elderly Medicare and Medicaid patients with colon cancer. Arch Intern Med. 2008 Mar 10;168(5):521-9.

- **METHODS:** Medicaid and Medicare administrative data were merged with the Michigan Tumor Registry to extract a sample of patients who had resection for a first primary colon tumor diagnosed between January 1, 1997, and December 31, 2000 (n = 4765).
- **RESULTS:** Relative to Medicare patients, Medicaid patients were less likely to initiate chemotherapy (odds ratio, 0.50; 95% confidence interval, 0.39-0.65) or complete chemotherapy.

Adjusted Logistic Regression Models of Adjuvant Chemotherapy Initiation, Chemotherapy Completion, and Evaluation by a Specialist for Patients With TNM-Staged Disease

Bradley, C. J. et al. Arch Intern Med 2008;168:521-529.

Table 5. Adjusted Logistic Regression Models of Adjuvant Chemotherapy Initiation, Chemotherapy Completion, and Evaluation by a Specialist for Patients With TNM-Staged Disease*

Explanatory Variables	Odds Ratio (95% Confidence Interval)		
	Adjuvant Chemotherapy Initiation (n=2873)	Chemotherapy Completion (n=188)	Oncology Evaluation (n=188)
Medicare only	1 (Reference)	1 (Reference)	1 (Reference)
Medicaid	0.49 (0.30-0.74) [‡]	0.52 (0.29-1.05)	1.01 (0.41-1.69)
Age, y	1 (Reference)	1 (Reference)	1 (Reference)
60-69	0.73 (0.53-0.99) [‡]	0.79 (0.46-1.35)	1.05 (0.61-1.80)
70-74	0.40 (0.28-0.56) [‡]	0.74 (0.46-1.20)	1.13 (0.67-1.92)
75-79	0.11 (0.08-0.15) [‡]	0.50 (0.29-0.82) [‡]	0.90 (0.59-1.36)
Race	1 (Reference)	1 (Reference)	1 (Reference)
White	0.77 (0.53-1.11)	0.89 (0.46-1.76)	0.88 (0.53-1.45)
African American or other	1 (Reference)	1 (Reference)	1 (Reference)
Male	0.50 (0.29-1.15)	0.71 (0.30-1.52)	1 (Reference)
Female	1 (Reference)	1 (Reference)	1 (Reference)
Comorbidity score	1 (Reference)	1 (Reference)	1 (Reference)
0	0.66 (0.50-0.86) [‡]	0.63 (0.50-1.20)	0.94 (0.66-1.33)
1	0.70 (0.50-0.96) [‡]	0.49 (0.27-0.86) [‡]	2.05 (1.29-3.17) [‡]
Hospital readmission	1 (Reference)	1 (Reference)	1 (Reference)
No	1.15 (0.79-1.65)	1.38 (0.70-2.49)	1.54 (0.93-2.56)
Yes	1 (Reference)	1 (Reference)	1 (Reference)
Teaching status	1 (Reference)	1 (Reference)	1 (Reference)
Yes	0.65 (0.51-0.83) [‡]	1.40 (0.90-2.04)	1.52 (0.87-1.70)
No	1 (Reference)	1 (Reference)	1 (Reference)
Census tract median annual income, \$	1 (Reference)	1 (Reference)	1 (Reference)
< 25,000	0.67 (0.58-1.30)	1.42 (0.70-2.89)	0.76 (0.41-1.24)
25,000-49,000	0.87 (0.70-1.08)	1.08 (0.98-1.19)	0.80 (0.64-0.99)
50,000-99,000	1.20 (0.89-1.73)	1.46 (0.80-2.66)	0.63 (0.48-1.41)
≥ 100,000	1 (Reference)	1 (Reference)	1 (Reference)
Urban or rural	1 (Reference)	1 (Reference)	1 (Reference)
Metropolitan	1 (Reference)	1 (Reference)	1 (Reference)
Rural, adjacent to metropolitan (bordered rural)	0.59 (0.40-0.81)	1.08 (0.57-1.92)	0.95 (0.51-1.81)
Rural, not adjacent to metropolitan (isolated rural)	0.57 (0.25-1.29)	1.00 (0.15-6.50)	0.15 (0.07-0.32) [‡]
Urban, not adjacent to metropolitan	0.60 (0.49-1.20)	0.64 (0.28-1.45)	0.67 (0.16-4.45) [‡]
Urban, adjacent to metropolitan	1.28 (0.83-1.90)	0.65 (0.30-1.20)	0.60 (0.23-1.09)
TNM stage	1 (Reference)	1 (Reference)	1 (Reference)
0	0.69 (0.09-0.04) [‡]	0.17 (0.08-0.36) [‡]	0.72 (0.41-1.26)
I	0.44 (0.11-1.64) [‡]	0.44 (0.17-1.04) [‡]	1.01 (0.73-1.29)
II	1 (Reference)	1 (Reference)	1 (Reference)
III	0.65 (0.29-0.50) [‡]	NA	0.79 (0.41-1.52)

Abbreviation: NA, not applicable.
[‡] All variables shown in the table are included in the adjusted logistic regression.
[‡] Statistically significant at P < .05.

Medicaid Data in North Carolina

Source: Foley KL, et al. Adjuvant chemotherapy among Medicaid-enrolled patients diagnosed with nonmetastatic colon cancer. Am J Clin Oncol. 2011 Apr;34:120-4.

TABLE 3. Percentage of People Receiving Chemotherapy Diagnosed With Regional Colon Cancer by Age and Comorbidity Status

Comorbidity	Age		
	<65	65-74	75+
Yes (1+)	89 (62.9%)	99 (55.6%)	209 (18.7%)
No (0)	117 (75.2%)	78 (47.4%)	100 (18.0%)

Medicaid Data Validation – high disenrollment after Dx

- A study by Ramsey SD, et al. Med care. 2008 Jan;46(1):49-57.
- **METHODS:** Medicaid enrollment and claims files were linked to the Washington State Cancer Registry to identify all Medicaid enrollees with breast, cervical, lung, colorectal, and prostate cancer between 1997 and 2002.
- **RESULTS:** We identified 5009 newly diagnosed cancer patients covered by Medicaid, approximately 13% of the total cases diagnosed in subjects less than 65 years of age in Washington State. The majority, 2866 (57%), enrolled in Medicaid around the time of diagnosis; the remainder had been enrolled at least 3 months before diagnosis. Persons enrolled at diagnosis had later-stage cancer; those enrolled before diagnosis had more noncancer comorbidities. Overall, 18% had disenrolled by 6 months after diagnosis; 34% by 1 year; and 54% by 2 years.
- **CONCLUSIONS:** Medicaid patients with cancer in Washington State experience a high rate of disenrollment within 1 year after diagnosis. Further research is needed to determine whether disenrollment compromises initial therapy or follow-up care."

Medicaid Data Validation – high rate of incomplete treatment

- Another study by Ramsey SD, et al. Breast J. 2010 Jan-Feb;16(1):20-7.
- Aims: "We evaluate radiation treatment completion rates for Medicaid enrollees aged 18-64 diagnosed with breast cancer, using data from the Washington State Cancer Registry linked to Medicaid enrollment and claims records, we identified Medicaid enrollees diagnosed with breast cancer from 1997 to 2003 who received BCS. Among the 402 women who met inclusion criteria, 105 (26%) did not receive any radiation."
- Results: ...
- Among those who received at least one radiation treatment, 65 (22%) failed to complete therapy and 71 (24%) patients had at least one 5 to 30 day gap in treatment.
- Conclusions:...
- A substantial proportion of Medicaid-insured women who are eligible for radiation therapy following BCS either fail to receive any treatment, experience significant interruptions during therapy, or do not complete a minimum course of treatment.
- - this was similar pattern/finding to that of SEER and/or Medicare data -

Medicaid Data Validation (in Texas)

- 1. Coker AL, Eggleston KS, Du XL, Ramondetta L. Ethnic Disparities in Cervical Cancer Survival among Medicare Eligible Women in a Multi-Ethnic Population. International Journal of Gynecological Cancer. 2009;19(1):13-20.
- 2. Coker AL, Eggleston KS, Du XL, Anderson M. Does race affect cancer survival for cervical cancer among women with a Medicaid coverage? Texas Cancer Registry - Medicaid claims linked sample, 1999-2001. under review.

Table. Treatment Data in Women with Cervical Cancer in 1999-2001: Texas Cancer Registry Vs. Medicaid Claims Data

Individual Level Characteristics	Non-Hispanic White N=258	Non-Hispanic Black N = 151	Hispanic N= 340
Treatment: Medicaid^{††}			
No treatment (REF)	100 (38.76)	69 (45.70)	122 (35.88)
Any treatment	158 (61.24)	82 (54.30) (p=0.17)	218 (64.12) (p=0.47)
Type of Treatment: Medicaid^{††}			
No Treatment (REF)	100 (38.76)	69 (45.70)	122 (35.88)
Cone or Hysterectomy	51 (19.77)	19 (13.24) (p=0.14)	93 (27.35) (p=0.05)
Chemotherapy	81 (31.40)	37 (24.50) (p=0.14)	111 (32.65) (p=0.75)
Radiation	124 (48.06)	70 (46.36) (p=0.74)	159 (46.76) (p=0.75)
Treatment: TCR^{††}			
No treatment (REF)	47 (18.22)	22 (14.57)	48 (14.12)
Any treatment	211 (81.78)	129 (85.43) (p=0.34)	292 (85.88) (p=0.17)
Type of Treatment: TCR^{††}			
No Treatment (REF)	47 (18.22)	22 (14.57)	48 (14.12)
Cone or hysterectomy	98 (37.98)	40 (26.49) (p=0.02)	109 (40.00) (p=0.82)
Chemotherapy	99 (38.37)	49 (32.45) (p=0.23)	139 (40.88) (p=0.54)
Radiation	141 (54.65)	96 (63.58) (p=0.08)	205 (60.29) (p=0.17)

* Non-Hispanic White is the reference group for Non-Hispanic Black and Hispanics;
^{††} p for trend; †† Treatment documented up to 1 year after cervical cancer diagnosis.

Potentials/Advantages of Medicaid Data

- Large # of community-based cases for low-income populations by states and across the US states.
 e.g. 'In any given month there are > 3 million clients enrolled in the Texas Medicaid Program' (source: Bowcock@TCR);
- Study <65 as well as >=65 yrs of low-income populations (children, pregnant women, elderly, & people with disability);
- Help identify additional CA cases (that may be missed by tumor registry);
 e.g., (source: Bowcock@TCR):
 45,226 Medicaid records contained a malignant cancer-related ICD-9 diagnosis code, but there was no record in TCR data.
 - 4,086 prostate (9.0% of the non-linked Medicaid records)
 - 5,853 female breast (12.9%)
 - 3,211 colorectal (7.1%)
 - 756 melanoma (1.7%)

Potentials/Advantages of Medicaid Data (2)

- Enhance the treatment data:
 - that were not recorded in tumor registry;
 e.g., radiation therapy for breast ca (source: Bowcock@TCR): Reported to TCR: 2,361; and not reported 1,183 (33%).
 - incremental value of Medicaid claims on comorbidity and treatment over other data sources (but small over that from Medicare claims).

source: Bradley CJ, et al. Incremental value of using Medicaid claim files to study comorbid conditions and treatments in dually eligible beneficiaries. Med Care. 2010 Jan;48(1):79-84.

Comorbidity	Non-Hispanic			Hispanic			White		
	Medicaid only	Medicaid and Medicare	% Non-Hispanic Medicaid and Medicare (95% CI, %)	Medicaid only	Medicaid and Medicare	% Hispanic Medicaid and Medicare (95% CI, %)	Medicaid only	Medicaid and Medicare	% White Medicaid and Medicare (95% CI, %)
Heart disease	31.9 (24.1)	3.8 (2.6)	31.6 (24.4)	30.9 (24.9)	3.50 (2.3)	30.9 (24.9)	2.8 (1.3 - 3.4)		
Stroke	1.0 (1.7)	0.0 (0.2)	1.0 (1.7)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.2 (0.1 - 0.3)		
Chronic lung disease	20.0 (16.7)	0.0 (0.0)	20.0 (16.7)	20.0 (16.7)	0.0 (0.0)	20.0 (16.7)	0.0 (0.0)		
Diabetes	4.0 (3.0)	0.0 (0.0)	4.0 (3.0)	4.0 (3.0)	0.0 (0.0)	4.0 (3.0)	0.0 (0.0)		
Arthritis	1.0 (1.3)	0.0 (0.0)	1.0 (1.3)	1.0 (1.3)	0.0 (0.0)	1.0 (1.3)	0.0 (0.0)		
Chronic kidney disease	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)		
Alcohol use disorder	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)		
Depression	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)		
Substance use disorder	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)		
Chronic pain	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)		
Obesity	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)		
Chronic liver disease	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)		
Chronic obstructive pulmonary disease	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)		
Chronic urinary tract infection	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)		
Chronic sinusitis	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)		
Chronic ear, nose, and throat disease	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)		
Chronic back pain	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)		
Chronic neck pain	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)		
Chronic headache	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)		
Chronic fatigue	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)		
Chronic dizziness	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)		
Chronic vertigo	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)		
Chronic tinnitus	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)		
Chronic insomnia	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)		
Chronic anxiety	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)		
Chronic depression	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)		
Chronic bipolar disorder	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)		
Chronic schizophrenia	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)		
Chronic personality disorder	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)		
Chronic eating disorder	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)		
Chronic self-harm	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)		
Chronic suicidal thoughts	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)		
Chronic suicidal behavior	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)		
Chronic suicidal ideation	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)		
Chronic suicidal communication	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)		
Chronic suicidal planning	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)		
Chronic suicidal attempt	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)		
Chronic suicidal ideation with intent	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)		
Chronic suicidal ideation without intent	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)		
Chronic suicidal ideation with and without intent	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)		
Chronic suicidal ideation with suicidal ideation	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)		
Chronic suicidal ideation with suicidal ideation and suicidal ideation	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)		
Chronic suicidal ideation with suicidal ideation and suicidal ideation and suicidal ideation	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)		

Potentials/Advantages of Medicaid Data (3)

- Provide other important treatment information (not covered by other programs), e.g. hormone therapy for breast ca (not covered by Medicare, before Part D in 2006).
 e.g., McLaughlin JM, et al. Patient and provider determinants associated with the prescription of adjuvant hormonal therapies following a diagnosis of breast cancer in Medicaid-enrolled patients. J Natl Med Assoc. 2009 Nov;101(11):1112-8.

Therapy	Number of Patients (%) by Year of Drug Index					
	2000	2001	2002	2003	2004	T total
Tamoxifen (only)	35 (94.6)	132 (82.5)	105 (59.7)	26 (41.9)	9 (52.9)	307 (67.9)
Aromatase inhibitor (ever)	2 (5.4)	28 (17.5)	71 (40.3)	36 (58.1)	8 (47.1)	145 (32.1)

Potentials/Advantages of Medicaid Data (4)

- Valuable to evaluate the CDC's National Breast and Cervical Cancer Early Detection Program (NBCCEDP): uninsured and underinsured women at or below 250% of federal poverty level; ages 18–64 for cervical screening; ages 40–64 for breast screening;
 - Chien LN, et al. Treating cervical cancer: Breast and Cervical Cancer Prevention and Treatment Act patients. *Am J Obstet Gynecol.* 2011 Jun;204(6):533.e1-8.
CONCLUSION: Treatment patterns among Georgia Medicaid cases appear appropriate to stage but 18% with invasive cervical cancer received no cancer treatment, although Medicaid enrolled.
 - Texas Breast Cancer Screening, Early Detection and Treatment Program Outcomes Study (PI: Dr. Begley at SPH) funded by CPRIT, to understand the benefits and failings in the Texas Breast and Cervical Cancer Services (BCCS) program and develop innovative strategies for program improvement and expansion to a larger proportion of the low income population.

Potentials/Advantages of Medicaid Data (5)

- Other advantages
 - Data are inexpensive to obtain (relative to RCTs or cohort studies, etc).
 - Unobtrusive nature: no patient contact, patient consent not sought (only needs to sign Data User Agreement), and quick IRB review.
 - Flexibility of Study Design: cohort, case-control, etc.
 - Potential for Multiple Projects.
 - Enhanced by linking with tumor registry data, census data, and others.
 - Can provide valuable information on comorbidity (generate comorbidity index/scores);
 - Help identify 'most vulnerable/high risk pop.' for emergency preparedness (City of Houston study).
 - Used to study many other health events/conditions such as childhood asthma and HPV vaccine etc.

Pitfalls of Medicaid Data

- Not initially designed for research, but for administrative purposes, so that incomplete or inaccurately coded bills, and down/up-coding posed potential problems;
- Unclear how the 'enrolled' represent the 'eligible' for Medicaid? – selection bias & generalizability ?
- 'Population-at-risk' vs. 'high-risk population': large number of cases enrolled with 'health events' of study interest. – different groups of pop.?
e.g., source: Schrag D, et al. Linking tumor registry and Medicaid claims to evaluate cancer care delivery. *Health Care Financ Rev.* 2009;30(4):61-73.

Medicaid Enrollment Status at month of diagnosis	Number of Patients (N=6,800)	Percent of Cohort
YES, Enrolled during month of diagnosis	5,351	79%
NO, First enrolled after month of diagnosis ¹	1,247	18%
NO, Enrolled prior to diagnosis but not in month of diagnosis	202	3%
Duration of Medicaid Enrollment prior to diagnosis		
<=12 months	3,323	49%
>12 but <=6 months	1,450	21%
<6 months	578	9%
Not enrolled until AFTER diagnosis	1,247	18%
Enrolled prior to diagnosis, but not during month of diagnosis	202	3%

¹Of these 714 or 57%, enrolled in Medicaid within 3 months of diagnosis.

Pitfalls of Medicaid Data (cont'd-2)

- The Medicaid program is essentially an insurance program with federal oversight and state administration. Since states have latitude in program coverage decisions, it may be difficult to complete some analyses across states and/or across years. Eligibility rules and program benefits can vary across states and across years.
 - Data from CMS
 - Data from a state
- Q: Are non-CMS Medicaid data files available? (from CMS-ResDac website):
 - You may be able to obtain more recent claims data directly from a state. Files available from states may be in the form of MSIS data tapes or other state specific file formats. MSIS files also are not in the format of final action claims, meaning the researcher will need to apply algorithms or otherwise determine which claims were finally paid and settled. It may be more difficult to do comparisons across years and states since the files will not have a uniform format. MSIS files obtained from states will not have gone through CMS validation and editing procedures.

Pitfalls of Medicaid Data (cont'd-3)

- Medicaid enrollments: – non-continuous and on/off multiple times? Same or different IDs? (versus Medicare program – 'life long')
- A 'high' rate of disenrollment within 1 yr after diagnosis → challenge for FU care study, recurrence & late effects of Ca/treatments
- Lack of external validation of Medicaid information (accuracy and completeness) on many data items (such as chemotherapy agents, hormone therapy etc.)

Pitfalls of Medicaid Data (cont'd-4)

- Other limitations:
 - No reliable information on defining cancer stage;
 - Incomplete data for Medicaid enrollees who belonged to HMOs, other insurance such as employer's health plan and those who received care in Veterans Affairs hospitals.
- Q: Can managed care services be studied with the CMS Medicaid claims data files? (from CMS-ResDac website)
 - There are 2 approaches to identifying Medicaid beneficiaries who were covered under a managed care plan.
 - The Personal Summary File indicates if a beneficiary was in a managed care plan with coverage purchased by the state, and this information is indicated for each month.
 - It may be necessary for other types of plans to determine if the managed care is medical, dental, transportation or some other form. This information can be obtained by examining claims in the Other Therapeutic Claims file. Claims with a Type of Service = '20' (premium payment) indicate beneficiaries covered by a managed care plan.

Thank you for Attention!

Questions/Comments